

O Método MCMC a Dois Estágio Aplicado em Problemas de Transporte Linear

Thiago Jordem Pereira¹

Depto de Ciências Exatas, Biológicas e da Terra, INFES, UFF, Santo Antônio de Pádua, RJ

Helio Pedro Amaral Souto²

Depto de Modelagem Computacional, Instituto Politécnico, UERJ, Nova Friburgo, RJ

Resumo. O objetivo deste trabalho é o estudo da quantificação de incertezas em problemas de escoamentos de traçadores em meios porosos heterogêneos. Na parametrização dos campos de permeabilidades utiliza-se o método SSCGI, onde o campo aleatório Gaussiano é expresso em termos de uma função de covariância do tipo exponencial. Emprega-se uma abordagem Bayesiana para a atualização dos campos de permeabilidades, baseando-se em um conjunto de medições observadas da concentração do traçador. Na resolução do problema inverso, recorre-se ao método MCMC a dois estágios que utiliza o algoritmo de Metropolis-Hastings baseado em passeios aleatórios do tipo autoregressivo.

Palavras-chave. Escoamentos em meios porosos, Função de covariância exponencial, Incertezas, Métodos MCMC a dois estágio, Traçador.

1 Introdução

O modelo matemático que descreve o fenômeno do escoamento de traçadores em meios porosos heterogêneos constitui uma ferramenta essencial para o processo de tomada de decisão voltada para a produção de óleo em reservatórios de petróleo e o monitoramento do transporte de contaminantes em formações subterrâneas.

As soluções das equações diferenciais parciais que constituem o modelo do problema de escoamento, são fortemente influenciada pelas variações significativas das propriedades do meio poroso, tais como a porosidade e a permeabilidade, que são responsáveis pela introdução das incertezas contidas no modelo matemático. Com o auxílio da metodologia Bayesiana é possível reduzir de forma significativa as incertezas relacionadas com a caracterização das propriedades do meio poroso e, conseqüentemente, melhorar a confiabilidade das previsões de interesse [4, 6, 8]. Nesta metodologia, o Teorema de Bayes é utilizado como um mecanismo de atualização dos parâmetros físicos do modelo geológico, levando em consideração a distribuição *a priori* e a distribuição amostral dos dados (função de verosimilhança), obtendo a distribuição *a posteriori* do parâmetro de interesse [4, 6, 8].

As previsões são feitas a partir da resolução deste problema inverso complexo, onde pode-se utilizar os métodos estocásticos MCMC, abreviatura do inglês *Markov Chain*

¹tjordem@gmail.com

²helio@iprj.uerj.br

Monte Carlo (Monte Carlo via Cadeias de Markov) [4, 6, 8]. Estes métodos baseiam-se na simulação de cadeias de Markov ergóticas cuja distribuição estacionária é a distribuição de interesse, isto é, a distribuição de probabilidade *a posteriori*. A utilização dos métodos MCMC em problemas de escoamento de fluido em meios porosos exige um grande esforço computacional e leva a uma baixa taxa de aceitação das amostras propostas do parâmetro de interesse. Desta forma, se torna necessário à proposta de empregar-se os métodos MCMC a dois estágios, que utiliza um modelo matemático, na escala de campo (malha grossa), para descrever o escoamento do traçador em um meio poroso [4, 6, 8].

Portanto, o objetivo deste trabalho é estudar da quantificação de incertezas em problemas inversos que envolve escoamentos de traçadores em meios porosos heterogêneos e propor atualização dos campos de permeabilidades, baseando-se no conhecimento prévio dos valores da concentração do traçador em alguns pontos específicos do meio poroso.

2 Modelagem Matemática nas Escalas Fina e Grossa

O sistema de equações diferenciais parciais, na escala fina, usado na modelagem do fenômeno do escoamento de traçadores em um meio poroso heterogêneo é constituído pelas lei de Darcy juntamente com a hipótese de incompressibilidade do fluido [4]

$$\mathbf{u} = -\frac{k(\mathbf{x})}{\mu} \nabla p, \quad \nabla \cdot \mathbf{u} = 0, \quad (1)$$

e pelo transporte advectivo linear da concentração do traçador

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \nabla c = 0, \quad (2)$$

onde $\mathbf{x} = (x, y)$ é o vetor posição, $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ e $p = p(\mathbf{x}, t)$ são, respectivamente, as velocidades de Darcy e a pressão da mistura fluida, μ representa a viscosidade da mistura, $c = c(\mathbf{x}, t)$ a concentração do fluido marcado e $k(\mathbf{x})$ é o campo escalar de permeabilidades.

Para a solução deste sistema, adota-se um domínio retangular $\Omega = [0, L_x] \times [0, L_y]$, com contorno $\partial\Omega$ e, um intervalo de tempo $0 \leq t \leq T$, com as condições de contorno dadas por $\mathbf{u} \cdot \mathbf{n} = -q$ em $x = 0$; $p = 0$ em $x = L_x$ e $\mathbf{u} \cdot \mathbf{n} = 0$ em $y = 0$ e L_y , onde \mathbf{n} é o vetor normal unitário exterior a $\partial\Omega$ e q é a taxa de injeção do traçador [4]. A condição inicial é $c(\mathbf{x}, 0) = 1$ se $x \leq 0$ (concentração do fluido injetado) ou $c(\mathbf{x}, 0) = 0$ se $x > 0$ (concentração do traçador no fluido residente) [4].

Na derivação das equações na escala grossa, utiliza-se uma técnica de *upscaling* [4, 6] de modo que as mesmas características do escoamento na malha fina sejam reproduzidas na malha grossa. Desta forma, pode-se obter um apropriado modelo matemático na malha grossa do problema do escoamento do traçador no interior do meio poroso [4, 6].

Na resolução do problema de pressão-velocidade (1), para as malhas fina e grossa, as equações diferenciais parciais são discretizadas pelo método dos elementos finitos mistos, que são adequados para o cálculo acurado dos fluxos nos campos de permeabilidades heterogêneos [1]. Na resolução da equação do transporte do traçador (2) emprega-se o método lagrangeano localmente conservativo *Forward Integral Tracking* (FIT) [2].

3 Parametrização das Incertezas

Na parametrização das incertezas contidas no modelo matemático do problema de escoamento do traçador, utiliza-se um modelo geológico probabilístico baseado em [7]

$$k(\mathbf{x}) = k_0 e^{\varphi Y(\mathbf{x})}, \quad (5)$$

onde $k_0 \in \mathbb{R}^+$ é uma constante, $\varphi \in \mathbb{R}^+$ controla o grau de heterogeneidade do meio e $Y(\mathbf{x})$ é uma campo aleatório Gaussiano caracterizado por sua média $\langle Y(\mathbf{x}) \rangle = 0$ e sua função de covariância do tipo exponencial [3]:

$$C_Y(\mathbf{x}_1, \mathbf{x}_2) = \sigma_Y^2 \exp\left(-\frac{|x_1 - x_2|}{L_x} - \frac{|y_1 - y_2|}{L_y}\right), \quad L_x > 0 \text{ e } L_y > 0, \quad (4)$$

onde $\mathbf{x}_1 = (x_1, y_1)$, $\mathbf{x}_2 = (x_2, y_2)$, L_x e L_y são os comprimentos de correlação nas direções x e y , respectivamente, e $\sigma_Y^2 = E(Y^2)$ é uma constante, onde $E(\bullet)$ é o operador esperança.

Os campos $Y(\mathbf{x})$ são gerados pelo método das Soma Sucessiva de Campos Gaussianos Independentes (SSCGI), que baseia-se na superposição hierárquica de malhas uniformes [3], isto é,

$$Y(\mathbf{x}) \equiv \underbrace{\sum_{n=0}^{\infty} \left(\sum_{m=1}^{b^n} \xi_{n_m}(\mathbf{x}) \right)}_{\xi_n(\mathbf{x})}, \quad (5)$$

onde $b^n \times b^n$, com $n = 0, 1, 2, \dots$, é o tamanho de cada elemento, $m = 1, 2, \dots, b^n$ é o número de malhas em cada nível n e, $\xi_{n_m}(\mathbf{x})$ e $\xi_n(\mathbf{x})$ são as variáveis aleatórias Gaussianas.

4 Método MCMC a Dois Estágios

No estudo da quantificação de incertezas para o problema inverso do transporte de um traçador em um meio poroso heterogêneo, emprega-se uma abordagem Bayesiana para a atualização dos campos de permeabilidades $k(\mathbf{x})$, baseando-se em um conjunto de medições observadas da concentração do traçador $O(c_o)$. Após observar $O(c_o)$, a informação contida em c_o fornecerá a verossimilhança dada por [4, 5]:

$$P(O(c_o)|k(\mathbf{x})) \propto \exp\left\{-\frac{\sum_{j=1}^{N_t} \sum_{i=1}^{N_s} [c_s(\mathbf{x}_i, t_j) - c_o(\mathbf{x}_i, t_j)]^2}{2\sigma_f^2}\right\}, \quad (6)$$

onde $c_o = c_o(\mathbf{x}_i, t_j)$ e $c_s(\mathbf{x}_i, t_j)$ são, respectivamente, as concentrações observada e simulada do traçador a partir do modelo matemático na escala fina (1)-(2), N_s é o número de sensores instalados nas localizações \mathbf{x}_i , N_t é o número de vezes em que a concentração é avaliada no tempo t_j e σ_f^2 é a precisão associada as medidas simuladas e observadas. Após a escolha de uma distribuição *a priori* $P(k(\mathbf{x}))$ [4] adequada, através do Teorema de Bayes a distribuição *a posteriori* $\pi(k(\mathbf{x})) = P(k(\mathbf{x})|O(c_o))$ expressa o novo conhecimento que se tem sobre os campos de permeabilidades [4, 5]:

$$\pi(k(\mathbf{x})) \propto P(O(c_o)|k(\mathbf{x})) P(k(\mathbf{x})). \quad (7)$$

O erro em malha fina é definido por [4, 5]:

$$\mathcal{E}_f = \sum_{j=1}^{N_t} \sum_{i=1}^{N_s} [c_s(\mathbf{x}_i, t_j) - c_o(\mathbf{x}_i, t_j)]^2. \quad (8)$$

De forma análoga, define-se função de verossimilhança, o erro e a distribuição *a posteriori* para o modelo matemático na escala grossa, onde utiliza-se $\bar{c}_s(\mathbf{x}_i, t_j)$ para a concentração simulada do traçador e σ_g^2 para a precisão associada ao modelo na escala grossa [4].

Os métodos MCMC a dois estágios têm sido frequentemente utilizados na exploração da complexa distribuição *a posteriori* do problema de escoamento de fluidos em meios heterogêneos [4,6,8]. Neste trabalho, as construções das cadeias de Markov $\{k^{(t)}\}$ são feitas mediante a utilização do algoritmo de Metropolis-Hastings baseado em passeio aleatório (*random walk*) do tipo autoregressivo [4]. Para cada iteração, propõe-se $Y(\mathbf{x})$ a partir do novo valor de $\xi_{n_m}(\mathbf{x})$ dado por

$$\xi_{n_m}(\mathbf{x}) = \sqrt{1 - h_\xi} \xi_{n_m}^{(t)}(\mathbf{x}) + \sqrt{h_\xi} z_\xi, \quad (9)$$

onde $\xi_{n_m}^{(t)}(\mathbf{x})$ representa o estado atual da variável aleatória Gaussiana, $n = 0, 1, 2, \dots$ e $m = 1, 2, \dots, b^n$, h_ξ é o parâmetro de perturbação e $z_\xi \sim \mathcal{N}(0, 1)$. O Algoritmo (1), descreve detalhadamente o método MCMC a dois estágios.

Algoritmo 1: Método MCMC a dois estágios.

- 1 **Passo:** Faça $t = 0$ e especifique um valor inicial $k^{(0)}$ tal que $\pi(k^{(0)}) > 0$.
- 2 **Passo:** Gere um candidato $\xi_{n_m} \sim q(\xi_{n_m} | \xi_{n_m}^{(t)})$. Então gere $k = k(\mathbf{x})$ dado por (3).
- 3 **Passo:** Resolva na malha grossa os problemas de pressão-velocidade e do transporte.
- 4 **Passo:** Gere uma variável aleatória $u \sim \mathcal{U}(0, 1)$ e teste se o campo proposto será aceito ou não:

$$k^{(t+1)} = \begin{cases} k & \text{(condicionalmente), se } u \leq \bar{\alpha}(k|k^{(t)}), \\ k^{(t)} & \text{(condicionalmente), caso contrário,} \end{cases} \quad (10)$$

onde

$$\bar{\alpha}(k|k^{(t)}) = \min \left[\frac{\bar{\pi}(k) q(k^{(t)}|k)}{\bar{\pi}(k^{(t)}) q(k|k^{(t)})}, 1 \right]. \quad (11)$$

Se o campo proposto k for rejeitado retorne ao **Passo 2**.

- 5 **Passo:** Resolva na malha fina os problemas de pressão-velocidade e do transporte.
- 6 **Passo:** Gere uma variável aleatória $u \sim \mathcal{U}(0, 1)$ e teste se o campo proposto será aceito ou não:

$$k^{(t+1)} = \begin{cases} k, & \text{se } u \leq \alpha(k|k^{(t)}), \\ k^{(t)}, & \text{caso contrário.} \end{cases} \quad (12)$$

onde

$$\alpha(k|k^{(t)}) = \min \left[\frac{\pi(k) Q(k^{(t)}|k)}{\pi(k^{(t)}) Q(k|k^{(t)})}, 1 \right]. \quad (13)$$

Se o campo proposto k for rejeitado, retorne ao **Passo 2**.

- 7 **Passo:** Faça $t = t + 1$, retorne ao **Passo 2** e continue o procedimento até atingir a convergência.
-

5 Resultados numéricos

Os resultados numéricos que serão apresentados tem como objetivo quantificar e reduzir as incertezas contidas no modelo matemático descrito na Seção 2. Em todas as simulações foi considerando um tempo máximo de 1.700 dias, em uma região física bidimensional com dimensões $L_x = 128$ m e $L_y = 128$ m. As malhas usadas nas escalas fina e grossa foram mantidas fixas e possuem 128×128 e 16×16 elementos, respectivamente. A taxa de injeção do traçador foi de 0,2 volumes porosos a cada ano. Na parametrização das incertezas utiliza-se $L_x^2 = L_y^2 = 20$ m e o fator que controla o grau de heterogeneidade do meio foi mantido fixo e igual a 0,4. Foram utilizados 6 sensores de monitoramento com as seguintes localizações: $S_1=(30,0, 25,0)$, $S_2=(40,0, 115,0)$, $S_3=(50,0, 80,0)$, $S_4=(55,0, 43,0)$, $S_5=(70,0, 10,0)$ e $S_6=(80,0, 98,0)$. As concentrações $c_s(\mathbf{x}_i, t_j)$ e $\bar{c}_s(\mathbf{x}_i, t_j)$ foram avaliadas a cada 1 dia de simulação, onde considerou-se várias relações entre as precisões associada aos modelos nas escalas fina e grossa: $\sigma_g^2 = \sigma_f^2$, $\sigma_g^2 = 2\sigma_f^2$, $\sigma_g^2 = 3\sigma_f^2$ e $\sigma_g^2 = 4\sigma_f^2$, com $\sigma_f^2 = 1,0$. O valor de h_ξ foi mantido fixo e igual a 0,0025.

No intuito de verificar se as cadeias de Markov geradas pelo método MCMC a dois estágios convergiram para a distribuição estacionária de interesse, utilizou-se a inspeção visual dos erros na malha fina [6], para os primeiros 620 campos de permeabilidades aceitos (veja a Figura 1).

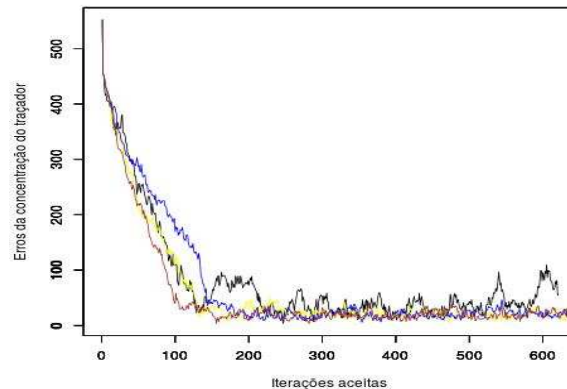


Figura 1: Erros na malha fina da concentração do traçador, com $\sigma_g^2 = \sigma_f^2$ (em preto), $\sigma_g^2 = 2\sigma_f^2$ (em amarelo), $\sigma_g^2 = 3\sigma_f^2$ (em azul) e $\sigma_g^2 = 4\sigma_f^2$ (em marrom).

Pode-se observar que as cadeias de Markov convergiram para a distribuição estacionária de interesse em todas as relações consideradas. Para os casos em que $\sigma_g^2 = 3\sigma_f^2$ e $\sigma_g^2 = 4\sigma_f^2$ os erros foram reduzidos de aproximadamente 552 para 14. Já para o caso $\sigma_g^2 = 2\sigma_f^2$, os erros foram reduzidos para valores mais elevados do que nos casos anteriores. A aceitação de amostras com valores menos elevados dos erros fornecer melhores valores médios da concentração do traçador medidos nos sensores. Chama-se a atenção para a maior variabilidade dos erros em malha fina em que $\sigma_g^2 = \sigma_f^2$. Neste caso, pode haver

uma diferença significativa entre os valores da concentração simulada e observada. Estes fatos podem ser claramente observados na Figura 2, onde compara-se os perfis médios de concentração simulada do traçador, tomando-se a média sobre um conjunto de 420 amostras aceitas dos campos de permeabilidades, após o período de aquecimento (*burn-in*) de 200 campos aceitos.

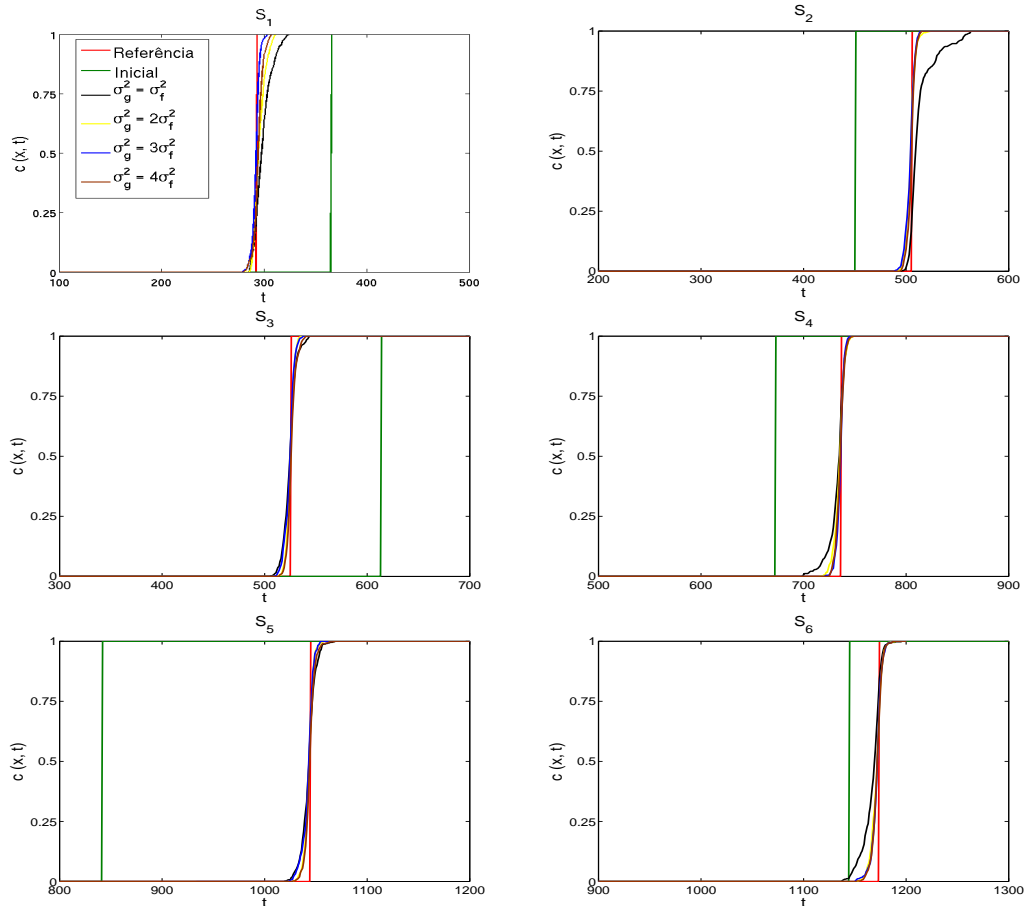


Figura 2: Valores da concentração do traçador em diferentes sensores.

Comparando as taxas de aceitação do método MCMC a dois estágios apresentadas pela Tabela 1, constata-se uma influência das relações entre as precisões σ_g^2 e σ_f^2 , onde as maiores taxas foram obtidas para $\sigma_g^2 = 3\sigma_f^2$ e $\sigma_g^2 = 4\sigma_f^2$. Entretanto, deve-se levar em consideração outros fatores que também influenciam na taxa de aceitação, tais como os valores das precisões na escala fina (σ_f^2) e refinamento das malhas na escala grossa.

6 Conclusão

Neste trabalho, foi estudado a quantificação de incertezas em problemas inversos envolvendo o escoamentos de traçadores em meios porosos heterogêneos empregando uma

Tabela 1: Taxas de aceitação para 620 primeiras amostras aceitas.

Relações	$\sigma_g^2 = \sigma_f^2$	$\sigma_g^2 = 2 \sigma_f^2$	$\sigma_g^2 = 3 \sigma_f^2$	$\sigma_g^2 = 4 \sigma_f^2$
Propostas	78.729	53.984	40.968	45.964
Aceitas: malha grossa	30.688	16.233	11.834	13.636
Aceitas: malha fina	620	620	620	620
Taxa de aceitação	2,02%	3,81%	5,23%	4,54%

abordagem Bayesiana para a atualização dos campos de permeabilidades, baseando na medição da concentração do traçador em diversos sensores instalados no meio poroso. Foram realizados estudos numéricos para um conjunto de realizações dos campos de permeabilidades, onde considerou-se diversas relações entre as precisões associada aos modelos nas escalas fina e grossa. Os resultados numéricos obtidos indicam que a metodologia utilizada foi capaz de quantificar e reduzir as incerteza contidas no modelo matemático que descreve o problema do transporte de traçadores.

Referências

- [1] J. Aquino, A. S. Francisco, F. Pereira, and H. P. A. Souto. Numerical simulation of transient water infiltration in heterogeneous soils combining central schemes and mixed finite elements, *Comm. Num. Met. Eng.*, 23: 491–505, 2007.
- [2] J. Aquino, A. S. Francisco, F. Pereira, and H. P. A. Souto. A hybrid method for the simulation of radionuclide contaminant plumes in heterogeneous unsaturated formations, *Prog. Nucl. Ener.*, 53: 1159–1166, 2011.
- [3] M. R. Borges, F. Pereira, and H. P. A. Souto. Efficient generation of multi-scale random fields: a hierarchical approach, *Int. J. Num. Met. Bio. Eng.*, 26: 176–189, 2010.
- [4] M. R. Borges, T. J. Pereira, and H. P. A. Souto. Uncertainty quantification for tracer transport in heterogeneous porous media using two-stage MCMC method and autoregressive instrumental distribution., *IRECHE (Testo Stam.)*, 5: 400-408, 2013.
- [5] C. Douglas, Y. Efendiev, R. Ewing, V. Ginting, R. Lazarov, Dynamic Data Drive Simulations in Stochastic Environments, *Computing*, **77** (2006), 321–333.
- [6] Y. Efendiev, T. Hou, and W. Lou. Preconditioning Markov Chain Monte Carlo simulations using coarse-scale models, *SIAM Jour. Sci. Comp.*, 2 (2): 776–803, 2006.
- [7] F. Furtado and F. Pereira. Crossover from nonlinearity controlled to heterogeneity controlled mixing in two-phase Porous media flows, *Comp. Geo.*, 7 (2): 115–135, 2003.
- [8] G. Victor, P. Felipe, P. Michael, and W. Shaochang. Application of the two-stage Markov Chain Monte Carlo method for characterization of fractured reservoirs using a surrogate flow model. *Comp. Geo.*, 15 (4): 691–707, 2011.